

A lexikai kohézió vizsgálata a fordításkutatásban egy korpusznyelvészeti eszközzel végzett kutatás tükrében

Péché Olívia

Eötvös Lóránd Tudományegyetem, Bölcsészettudományi Kar
Nyelvtudományi Doktori Iskola, Fordítástudomány Program
Múzeum krt. 4/F, 1088 Budapest, Hungary
pech_oli@t-online.hu

Kivonat A lexikai kohéziós vizsgálat a szöveg jelentésbeli összefüggés-hálózatát kutató, az önálló mondatokból kialakuló szöveg szerveződésének jellemzőit vizsgáló szövegsemantikai megközelítés részterületeként a szöveg szempontjából releváns lexikai elemek ismétlődését tárja fel. Mivel a különböző nyelvek különbözőképpen tolerálják a lexikai kohéziós eszközök alkalmazását, a fordításra feltételezésem szerint Catford terminusával élve shift, azaz eltolódás jellemző ebben a tekintetben is. Célkitűzésem annak feltérképezése, hogy kimutatható-e adott nyelvpárok között a fordítás következtében megváltozó kohéziós eszközhasználat esetében valamilyen törvényszerűség a változás jellegét illetően. Bemutatom azt az általam végzett, a számítógépes korpuszelemzés előnyeit kihasználó lexikai kohéziós vizsgálatot, amelynek során az autentikus szövegre jellemző lexikai kohéziós elemek számát és eloszlásának arányát hasonlítom össze az adott szöveg fordításaként keletkezett célnyelvi változatra jellemző lexikai kohéziós mintázattal.

1 Bevezetés

Jelen dolgozatomat négy fő kérdés köré csoportosítom. Elsőként ismertetem a kohéziókutatás és fordításkutatás közös érdeklődésére számot tartó, a szakirodalomban a két diszciplína összekapcsolásával eddig tárgyalt kutatási területeket, illetve azt az új megközelítést, amelyet saját kutatásom valósít meg. Másodsorban vázolom az új megközelítés technikai háttérét biztosító számítógépes eszközökkel segített kutatás előnyeit, az empirikus szövegkutatás számára új távlatokat nyitó korpusznyelvészeti elemzés nyilvánvaló és rejtettebb lehetőségeit. Ezt követően bemutatom hipotézisemet és kérdésfelvetésem, amit végül a lexikai kohézió vizsgálatára általam kiválasztott elemzési módszer, és az az elemzés követ, amelyben egy korpusznyelvészeti eszköz segítségével az eredeti magyar szövegekre és német fordításaikra jellemző lexikai kohéziós eszközhasználatot hasonlítom össze.

2 Lexikai kohéziókutatás és fordítástudomány

2.1 A lexikai kohézió vizsgálata

A lexikai kohézió vizsgálata a szövegkutatás egyik meghatározó irányzata, melynek tárgya elsősorban írott szövegek leíró, funkcionális elemzése. A szövegközpontú szemlélet az 1960-as évek végétől vált népszerűvé, mivel a mondatokból szöveggé szerveződött nyelvi formák kapcsán felvetődő kérdések nem mindegyikét lehetett már a mondatgrammatika keretein belül megválaszolni (Tolcsvai Nagy 1994). A szöveget mindinkább egységes egésznek, a nyelvhasználat egységének tekintő felfogás szerint a szöveg és alkotómondatainak kapcsolatában a hangsúly már nem arra került, hogy a szöveg mondatokból áll, hanem inkább a szöveg mondatok által történő realizációjára (Halliday és Hasan 1976).

A szöveg jelentésbeli összefüggés-hálózatát kutató szövegsemantikai megközelítés a szöveg szempontjából releváns lexikai elemek ismétlődését vizsgálja. Meghatározásában „a kohézió a szöveg egy elemének a szemantikai viszonya egy másikhoz, amely elem az elsőnek az értelmezéséhez szükséges” (Halliday és Hasan 1976: 8). A szöveg kohéziós kapcsolatait feltáró tanulmányok a lexikai elemeket a szövegben betöltött szerepük szempontjából két csoportra osztják. A lexikai kohéziós kutatások középpontjában a centrális vagy releváns elemek állnak, amelyek a téma fő gondolatát viszik, míg a témához lazábban vagy egyáltalán nem kapcsolódó lexikai elemeket figyelmen kívül hagyják (Beaugrande és Dressler 2000).

2.2 Lexikai kohéziókutatás és fordítástudomány: eddigi közös kutatási területek és eredmények

A fordítástudomány, bár csak a múlt század második felétől vált önálló tudományterületté, viszonylag hamar nyitott más diszciplínák irányába. A fordítási feladatok természetéből, azaz a mondathatárokon túlmutató szövegegységek „átkódolásából” adódott a szöveg szintű vizsgálatok szükségessége, így többek között a forrásnyelvről célnyelvre történő fordítás során a kohéziós eszközhasználattal kapcsolatos megfigyelések feltárása. E megközelítések alapvetően két különböző fókuszpont körül kristályosodtak ki, képviselve egyrészt a leíró fordításkutatás deskriptív kategóriáit, másrészt a fordításoktatás preskriptív követelményrendszerét.

2.2.1 Leíró fordításkutatás

A szövegszerűség, másnéven textualitás kritériumainak meghatározására számos szövegnyelvészeti értekezés tesz kísérletet, melyek értékelése és egybevetése nem tartozik a jelen dolgozat célkitűzései közé. Tanulmányunk szempontjából kizárólag azon munkák említését tartjuk fontosnak, amelyek a szövegalkotással, és ezen belül is a lexikai kohézió vizsgálatával a fordítási tevékenység tükrében foglalkoznak.

Bár Heltai *A fordító és a nyelvi normák* című tanulmánysorozatában a „kohéziós eszközök használatát” a „szövegalkotási normák közé” sorolja (Heltai 2004: 417) mégis megállapítja, „hogya a szövegalkotási normák nem fogalmazhatók meg olyan egyértelműen és explicit módon, mint a morfológiai és szintaktikai normák”. Heltai

ennek ellenére a fordító feladatai közé sorolja, „a természetesnek vett „nyelvhelyességi normákon” kívül“ (Heltai 2005b: 169) „az optimális szöveg kohézió” biztosítását (Heltai 2005a: 36), illetve „a szöveg megformálása során elsősorban a [...] szövegszintű normákra, konvenciókra, illetve disztribúciós mintákra” való ügyelést, melynek részét képezi a megfelelő kohéziós eszközhasználat mondat és szövegszinten (Heltai 2005b: 169).

A fordított szövegek helyes kohéziójának funkcionális jelentőségét támasztja alá Klaudy kvázi-helyesség fogalma, melynek kapcsán arról beszél, hogy sokszor „a fordított magyar szövegek kohéziója csökken, az olvasó késve azonosítja a mondatok szerkezetét, nem világos a hangsúlyos és hangsúlytalan részek viszonya“ (Klaudy 1987: 7). A kvázi-helyes szövegek esetében a megszokott célnyelvi szövegalkotási normáktól való eltérés miatt „az olvasó számára sokkal nagyobb megterhelést jelent a mondanivaló kihámozása a szokatlan nyelvi formából” (Klaudy 1987: 7), illetve Klaudy megállapítása szerint az olvasók az „egész szöveget mégis elutasítják, mert nem felel meg a helyesen szerkesztett szövegről alkotott intuitív elképzeléseiknek“ (Klaudy 2004: 389).

Szabómihály a szlovákiai kétnyelvű célközönség számára készülő, többnyire nem önálló magyar nyelvű szövegproduktumokat vizsgálva sürgeti a fordítások értékeléséhez elengedhetetlen „objektív és egységes szempontrendszer” kialakítását, melynek szövegszinten a „retorikai cél; műfaj; a grammatikai és lexikai regiszter - szaknyelv; tematikus rend“ mellett a kohéziónak is meghatározó elemnek kell lennie (Szabómihály 2003: 58).

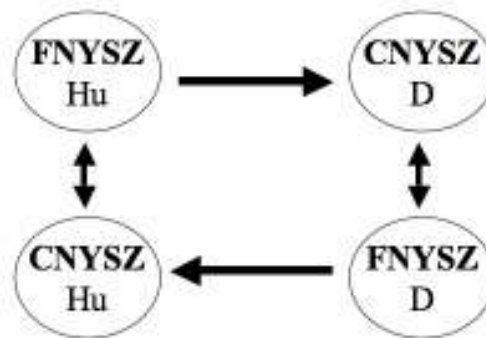
2.2.2 Fordításoktatás

A fordításoktatás számára a megfelelő lexikai kohéziós eszközök célnyelvi használatának helyes elsajátítása az előző pontban leírtakból következően alapcélkitűzésként jelenik meg. A fordítóképzés célja - mint ahogy azt Dróth tantárgyleírásában olvashatjuk - „az anyanyelvre fordított szöveg idegenszerűségek nélküli, a célnyelvi igényeknek és hagyományoknak megfelelő nyelvi megformálásához szükséges készségek kialakítása, fejlesztése“ (Dróth 2002: 95). A „megfelelő kohéziós eszközök alkalmazásának képessége“ megtalálható a fordítóképzés során folyamatosan finomítandó fordítási kompetencia részelemei között, s ez a terület nemcsak a készségek elsajátításakor kap figyelmet, hanem iránymutató szándékkal a Dróth féle ötszintű hibatipológia formatív értékelési szempontrendszerének kritériumai között is szerepel (Dróth 2002: 95). Ezen értékelési kritériumok viszont a fordítás eredményeként keletkezett célnyelvi szövegeket aszerint osztályozzák, hogy mennyiben felelnek meg a célnyelvi normák követelményrendszerének, mivel a „fordítások formatív értékelésekor már nem a forrásnyelvi szöveget vetjük egybe a célnyelvi szöveggel, hanem a célnyelvi szöveget elemezzük“ (Dróth 2002: 98).

2.3 A kutatásom során megvalósuló új megközelítés

Az előző két pontban vázolt szakirodalmi megközelítések alapvetésének igazságtartalmából kiindulva, miszerint a fordítások minőségének lényegi eleme a kohézió megléte, a saját kutatásom keretében éppen a fordítás során a kohéziós

elemek használatában bekövetkezett változások feltárására kívánok fókuszálni. A fordítástudományi és kohézió kutatási területek összekapcsolásakor általam kínált új megközelítés lényegét tehát az összehasonlító elemzés jelenti. E összehasonlító elemzés során első lépésben az autentikus magyar szövegekre és a fordításukként keletkezett német célnyelvi szövegekre jellemző kohéziós eszközhasználatot vizsgálom. Kutatásom későbbi szakaszában tervezem autentikus német forrásnyelvi szövegek és a fordításukként keletkezett magyar célnyelvi szövegek kohéziós eszközhasználatának összehasonlítását is, illetve a majdan rendelkezésre álló elemzett szövegek segítségével mind a magyar autentikus és fordított, mind a német autentikus és fordított szövegekre jellemző lexikai kohéziós elemek disztribúciójának egybevetését (1. ábra) (rövidítések: FNYSZ: forrásnyelvi szöveg, CNYSZ: célnyelvi szöveg, Hu: magyar, D: német).



1. ábra Összehasonlító elemzés

3 A számítógéppel segített elemzés

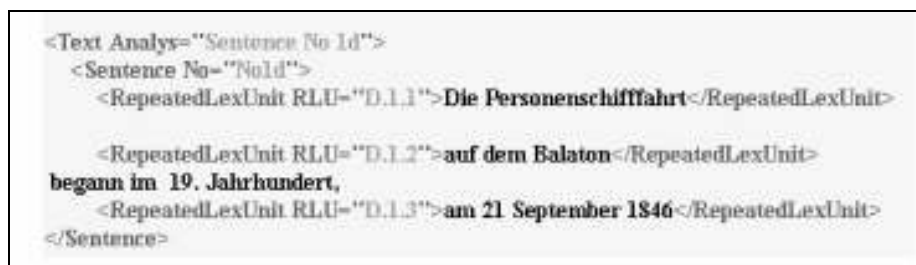
Az előbbieken felvázolt összehasonlító elemzésre természetesen nem véletlenül nem került sor a szövegkutatás korábbi periódusaiban, hiszen egy hagyományos eszközzel végzett lexikai kohézió kutatás keretein belül ez a fajta összehasonlítás igen körülményes és időigényes lenne. Mivel a hagyományos eszközzel végzett elemzés egy-egy szó kapcsolatait a szóból kiinduló, s az adott lexikai egység ismétléses kapcsolataihoz érkező vonalak segítségével jelöli (vö. Károly 2002: 80), egyszerűen belátható, hogy a szövegek, illetve mondatok párhuzamosítása milyen ember feletti feladatot jelentene.

Egy adott forrásnyelvi szöveg és annak célnyelvi fordításának mondat és szövegszintű egybevetése a napjainkban egyre népszerűbb korpusznyelvészeti módszer segítségével vált beláthatóvá és elérhetővé. Nem tartom ezért túlzottnak azt a megállapítást, hogy a fordításokban realizált lexikai kohéziós elemek egybevetése a forrásnyelvi szövegekre jellemző megvalósulásukkal alapvetően egy új technológia elterjedésének fejleményeként valósulhatott meg. A korpusznyelvészeti módszer kínálta lehetőségek új távlatokat nyitnak az alkalmazott nyelvészeti kutatások számára

mind az elemzések gördülékenyebbé tétele kapcsán, mind a lekérdezési lehetőségek terén.

3.1 A struktúrált kódolás

Az elemzésre kiválasztott szövegek feldolgozását, azaz a kohéziós eszközök bejelölését a CLaRK korpusznyelvészeti program segítségével végeztem el. A CLaRK program jellegzetessége, más hasonló nyelvészeti elemzésre írt programokhoz hasonlóan, hogy lehetővé teszi a számítógépes feldolgozásra szánt szövegek strukturált kódolását. Ez annyit jelent, hogy míg az eredeti szöveg integritása megmarad, addig annotációként többletinformációt, például a nyelvészeti elemzés eredményeit vihetjük be a szövegbe (Prószéky és Kis 1999). Ilyen szövegbe ágyazott többletinformáció a jelen elemzés esetében a mondathatárok bejelölése, illetve az ismétlődő lexikai elemek kiválasztása, és a kohézió típusának megjelenítése. Az eligazodásban és a szemléletességben segít, ha a CLaRK rendszerbe beolvasott eredeti szövegrészeket és az annotált információkat a képernyőn különböző színekkel jelenítjük meg (2. ábra).



```
<Text Analys="Sentence No Id">
  <Sentence No="NoId">
    <RepeatedLexUnit RLU="D.1.1">Die Personenschiffahrt</RepeatedLexUnit>
    <RepeatedLexUnit RLU="D.1.2">auf dem Balaton</RepeatedLexUnit>
    begann im 19. Jahrhundert,
    <RepeatedLexUnit RLU="D.1.3">am 21 September 1846</RepeatedLexUnit>
  </Sentence>
```

2. ábra Annotált szövegrészlet

3.2 Lekérdezési lehetőségek

Az elemzés gépi feldolgozásának köszönhetően ugyanabból az annotált kiindulási szövegből többféle lekérdezés alapján többféle összefüggés kimutatása is könnyen kivitelezhetővé válik. A lekérdezéseknél alapvetően két meghatározó szempontot módosíthatunk: egyrészt meg kell határoznunk, hogy mit, mely szöveg egységet kívánjuk az elemzés kiindulási pontjának tekinteni, másrészt hogy hogyan, milyen szempontrendszer alapján kívánjuk a kiválasztott szövegegységet vizsgálni.

Vizsgálatunk tárgya lehet egy szó, egy mondat, egy teljes szöveg, egy forrásnyelvi szöveg és annak célnyelvi fordítása, vagy éppen az összes elemzett szövegmondat és azok fordítása. Az elemzett szövegegységeknek a lekérdezések számára történő tetszőleges kiejelése teszi éppen racionálisan kivitelezhetővé az összehasonlító elemzést, mivel a forrásnyelvi és célnyelvi szövegek adekvát elemeinek párhuzamosítása (szó-, mondat- és szövegszinten) a gépi rendszer számára korántsem olyan lassú és bonyolult folyamat, mintha ezt a műveletet csupán a humán erőforrásokra támaszkodva kellene elvégezni.

A lekérdezések esetében meghatározható másik szempontot figyelembe véve a kutatás célkitűzéseinknek megfelelően a taggelt metainformációk segítségével különböző kritériumok alapján pásztázhatjuk végig a korpuszt. Lekérdezhetjük a bejelölt lexikai kohéziós kapcsolatok előfordulásának számát, az előfordulások százalékos gyakoriságát, vagy éppen a kohéziót biztosító lexikai elem szöveggörnyezetét, azaz elemezhetjük a kapott eredményeket akár statisztikák, akár konkordanciák segítségével is (Balaskó 2005).

3.3 A gépi feldolgozás előnyei

A fejezet bevezető részében már utaltam arra, hogy az általam végzett összehasonlító elemzés realitásának meghatározó eleme a gépi feldolgozás. A korpusznyelvészeti módszer alkalmazásának előnye azonban nem kizárólag az összehasonlító elemzés megvalósíthatóságában rejlik, hanem e feldolgozási keret több szempontból is igen kedvezően segíti a nyelvész elemző munkáját.

A gépi feldolgozás előnyeit élvezhetjük mind az elemzési fázisban, mind a lekérdezési szakban. A lexikai kohéziós elemek és kapcsolódásaik bejelölése során - a hagyományos módszerhez hasonlóan - az elemzés nem nélkülözheti a kutató nyelvész szellemi erőfeszítéseit, hiszen az ilyen típusú összefüggések annotálása még korántsem automatizált. Nagy könnyebbésséget jelent viszont a számítógéppel támogatott kutatás esetében, hogy a szöveg strukturálása nagymértékben igazodik a kutatási célhoz, ami viszont idő közben módosulhat vagy bővíthet is. A gépi feldolgozás előnyei közé tartozik a kutatás során, illetve az utólag felmerülő szempontrendszer integrálásának lehetősége. Bármikor módunk nyílik arra, hogy a bejelölt adatokat más összefüggések szempontjából csoportosítsuk, illetve a program lehetővé teszi az idő közben felmerülő igényeknek megfelelően az elemzési kategóriák bővítését, vagy változtatását. Ezek az egyedi beállítások és módosítások teszik még hatékonyabbá a munkát (Szirmai 2005).

A korpusznyelvészeti program segítségével végzett kutatások színes lekérdezési palettáját már az előző pontban ismertettem, itt már csak ismételten hangsúlyoznám, hogy a program segítségével ugyanabból az annotált szövegből többféle lekérdezésre is módunk nyílik. Ez igen hatékonyá teszi a munkát és lehetőséget teremt a részletek feltárására, az oksági összefüggések aprólékos megközelítésére. Fontos szempont továbbá, hogy a gép gyorsan és megbízhatóan, azaz száz százalékos pontossággal végzi el helyettünk a kívánt szempontrendszernek megfelelő kigyűjtéseket és számításokat. Figyelemre érdemes körülmény, hogy a lekérdezések eredményei is rögtön a számítógépen állnak rendelkezésünkre. Ez egyrészt a tárolás, másrészt az utólagos hasznosítás szempontjából lényeges, mivel így nem kell külön időt fordítanunk a kapott eredmények begépelésére.

4 Hipotézisem

Kutatásom kiindulópontjául az a feltételezésem szolgál, hogy a különböző nyelvek kohéziós eszközhasználata nem feleltethető meg száz százalékban egymásnak. Ha a különböző nyelvek valóban eltérő módon tolerálják a kohéziós elemek ismétlését,

akkor a fordításra Catford terminusával élve shift, azaz eltolódás lesz jellemző a használatuk tekintetében is (Catford 1965). E hipotézis validitását kívánom a jelen kutatás leendő eredményeivel alátámasztani, s ennek jegyében kezdtem meg a lexikai kohéziós elemek disztribúciójának vizsgálatát autentikus magyar és fordított német szövegek esetében.

Az előzőekből következően két kérdés vizsgálatát tartom érdekesnek, és megválaszolandónak: (1) Vajon az adott „A” forrásnyelvről egy adott „B” célnyelvre fordított szöveg milyen mértékben őrzi meg az eredeti forrásnyelvi szövegre jellemző kohéziós eszközhasználatot? (2) Kimutatható-e adott nyelvpárok között a fordítás következtében megváltozó kohéziós eszközhasználat esetében valamilyen törvényszerűség a változás jellegét illetően?

5 A kutatás bemutatása

5.1 Szövegválasztás

Kutatásomhoz internetes honlapokról töltöttem le úgymond hivatalosnak mondható leíró szövegeket, amelyek egy-egy intézményt, tájegységet, szervezetet stb. mutatnak be. Szövegválasztásomat három fő szempont határozta meg. (i) A kiválasztott szövegek stílusa hasonló a leíró jellegből fakadóan, ami kizárja az eltérő stílusú szövegekre jellemző, pragmatikai és stiláris okokra visszavezethető különbségeket a lexikai kohéziós elemek használatánál (Szikszainé Nagy 1999: 69). Ez a szempont azért lényeges, mert „a kohézió csak részben tekinthető a szöveg inherens (bennfoglalt) tulajdonságának. A kohézió foka / mértéke szövegtípusonként más és más lehet.” (Vass 2002: 9). E tényező figyelembe vételét a gyakorisági vizsgálatok, és azok eredményeinek összehasonlíthatósága indokolja. (ii) A vizsgált szövegek német fordítását minden bizonnyal profi fordítók készítették, hiszen a letöltésük forrását jelentő honlapok komoly marketing érdekeket szolgálnak. A fordítások jó minősége a garancia arra, hogy a vizsgált célnyelvi szövegek esetében a lexikai kohéziós elemek használatában tapasztalható eltérés nem a suta fordítói kompetenciával, hanem a nyelvek közötti, a kohéziós eszközök alkalmazásakor megfigyelhető eltéréssel magyarázható. (iii) A harmadik, ám nem elhanyagolható szempont, hogy a szövegek letöltésével a begépelés fáradtságát és ez által időt takarítottam meg.

5.2 A kutatás elméleti háttere

Kutatásom elméleti hátterül Károly Krisztina 2002-ben kidolgozott rendszere szolgál, amely a lexikai ismétlődés szövegszervező funkcióját vizsgálja. Károly taxonómiája a szöveg kohézióját biztosító lexikai egységként az olyan egységet definiálja, amelynek jelentése nem vezethető vissza alkotóelemeinek jelentésére.

Károly a kohéziós kapcsolatok között két alapvető típust különböztet meg: a lexikai és a szöveg-alapú relációkat. Az első csoportot további két alapvető alcsoportra osztva külön kategoriába sorolja azokat az eseteket, (i) amikor a lexikai egység változatlan formában ismétlődik (repetition), (ii) és azokat, amikor a lexikai

egységnek csak az információtartalma ismétlődik meg teljes egészében vagy részben egy vele szemantikai kapcsolatban álló lexikai egység formájában (synonymy, opposites, hyponymy, meronymy). (iii) Elemzési modelljében külön kategóriát képviselnek a szöveg-alapú relációk (instantial relations). Károly az ismétlés, a szinonima és az ellentét jellegét tekintve különbséget tesz egyszerű és derivált alakok között. Derivációról akkor beszél, ha a lexikai egység képzett szó formájában ismétlődik (Károly 2002: 95-107).

Categories	Examples	
LEXICAL RELATIONS	Simple	Derived
I. Same unit repetition		
1. repetition	<i>Writers</i> do not hesitate. <i>Writers</i> communicate thoughts.	<i>Writers</i> do not hesitate. They <i>write</i> down their thoughts.
II. Different unit repetition		
2. synonymy	It's healthy <i>to exercise</i> every day. After <i>working out</i> at the fitness center I feel like a new man.	They've <i>built</i> several bridges this year. This <i>construction</i> company is efficient.
3. opposites	She <i>hates</i> him. Well, you can't <i>love</i> everyone.	What she now feels is <i>hated</i> . But she used to <i>like</i> him.
4. hyponymy	Unfortunately, not all <i>birds</i> sing. But <i>cuckoos</i> do.	
5. meronymy	I can't use my <i>hands</i> . I've hurt my <i>fingers</i> .	
TEXT-BOUND RELATIONS		
6. instantial relations	The <i>teacher</i> laughed. I knew <i>Kitty</i> was a cheerful person (equivalence) I tell you the story of my uncle's <i>cat</i> . His name was <i>Alisy</i> . (naming)	

3. ábra Károly ismétlés-taxonómiája (Károly 2002:104)

Károly ismétlés-taxonómiája tehát a szövegfelszín szintaktikai-szemantikai hálójának vizsgálatát kilenc lexikai kohéziós kategória segítségével végzi (3. ábra). E kilenc kategóriát alkalmazom én is kutatásom során a szövegszerveződés szempontjából meghatározó szemantikai kapcsolatok feltárására a magyar forrásnyelvű és a fordításukként keletkezett német célnyelvi szövegek összehasonlításakor.

5.3 A korpusznyelvészeti eszközzel végzett elemzés bemutatása

Az elemzés kezdeti lépéseként a CLaRK rendszerbe beolvasott folyamatos szövegekben bejelöltem a mondatokhatárokat, illetve az egyes mondatokat sorszámmal láttam el (4. ábra). A mondathatárookra és mondatszámra vonatkozó többletinformáció az elemzést segítő program számára is értelmezhető adat, ami elengedhetetlen feltétele az eredeti - és a fordított dokumentum mondatszintű párhuzamosításának. A mondatsorszám (No1, No2...) után álló „d” betű jelzi, hogy itt a német szöveg első, második, stb. mondata áll, míg a magyar szöveg megfelelő mondatának sorszáma mellett egy „h” betű található.


```

<TextAnalysis>Sentence No">
<Sentence No="No1d">
  Die Personenschiffahrt auf dem Balaton begann im 19. Jahrhundert, am 21 September 1846.
</Sentence>
<Sentence No="No2d">
  An diesem Tage wurde das erste Dampfschiff "Kisfakadly" ins Wasser gelassen.
</Sentence>
<Sentence No="No3d">
  In den letzten 158 Jahren wurde unsere Firma zu einem bestimmenden Element des Tourismus am Balaton.
</Sentence>
<Sentence No="No4d">
  Unsere Gesellschaft ist die größte Schifffahrtsfirma auf dem See, ihre bekannteste Tätigkeit ist die Personenschiffahrt.
</Sentence>
<Sentence No="No5d">
  Der Großteil der Personenschiffahrt wird auf dem See von 26 Motorschiffen abgewickelt.
</Sentence>

```

4. ábra A mondatatók bejelölése és a mondatok sorszámozása

A következő lépésben az első mondatot elemeztem, azaz az első mondatban bejelöltem azokat a lexikai egységeket, amelyek a kohéziós kapcsolatok, a Károly-féle ismétlődési taxonómia szempontjából meghatározóak (2. ábra). Ezeket a későbbiekben ismétlődő lexikai egységeket (repeated lexical unit) hivatkozási számmal láttam el. A hivatkozási számok (D.1.1, D.1.2, D.1.3) utalnak arra, hogy a német szöveg első mondatának 1., 2. és 3. eleméről van szó, illetve hasonló logika szerint jelöltem meg a további mondatok ismétlődő lexikai elemeit is. A hivatkozási számok segítségével tudtam a ismétléses kapcsolatokat a kiindulási mondatok lexikai elemeivel összekapcsolni.

A következő lépésben az első mondatban kiválasztott lexikai egységek ismétlődéseit (lexical unit) kerestem meg az összes többi mondatban. Ezeket a kapcsolódás jellegét tekintve (cohesive relation) is elemeztem, és minden esetben bejelöltem a Károly taxonómia 9 kategóriájának (egyszerű és derivált ismétlés, egyszerű és derivált szinonima, egyszerű és derivált ellentét, meronima, hiponima és pillanatnyi relációk) megfelelő kohéziós kapcsolatot (5. ábra). A kohéziós kapcsolat meghatározást az a referenciális szám követi, amely visszaül az elemzés kiindulási mondatában hivatkozási számmal ellátott ismétlődő egységre.

Annak mintájára, ahogy az első mondat ismétlődő kapcsolatait meghatároztam, sorban elemeztem a szöveg összes többi mondatát is, és mindegyiknek megkerestem az ismétléses kapcsolatait a többi mondattal. Így jutottam el a teljes szöveg elemzéséhez, ami azt jelenti, hogy minden mondatot mindegyik mondattal összevettem.

```

<Sentence No="No1d">Unsere Gesellschaft ist die größte
  <LexicalUnit CohesiveRelation="InstantialRelation" reference="D.1.1">Schifffahrtsfirma</LexicalUnit>
  <LexicalUnit CohesiveRelation="SimpleSynonymy" reference="D.1.2">auf dem See</LexicalUnit>
  ihre bekannteste Tätigkeit ist
  <LexicalUnit CohesiveRelation="SimpleRepetition" reference="D.1.1">die Personenschiffahrt</LexicalUnit>
</Sentence>

```

5. ábra A kohéziós kapcsolatok bejelölése

5.4 Hol tart ma a kutatás?

Eddigi munkám során 3 teljes magyar szöveget és ezek német célnyelvi fordításait elemeztem. A szövegek mondatszáma 11 és 17 között váltakozik, ami az egyes szövegekhez kapcsolódó elemzések számát is meghatározza. Mivel egy elemzés legalább 2 szövegmondat lexikai elemeinek egybevetését jelenti, így a szövegenkénti elemzések száma 10 és 16 közötti mozog, eggyel kevesebb, mint az adott szövegre jellemző mondatszám. Pillanatnyilag így 82 elemzés áll a rendelkezésemre, azaz 82 kiindulási mondat esetében vizsgáltam az adott mondat lexikai elemei és a többi szövegmondat lexikai elemei között fennálló kohéziós kapcsolatot.

5.5 Lekérdezések

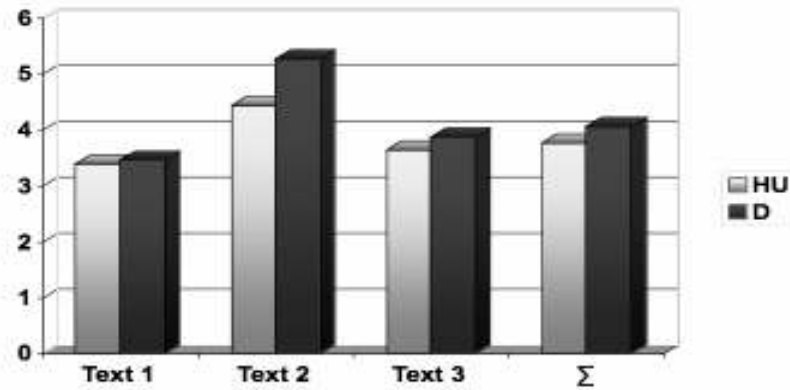
Tekintettel arra, hogy kutatásom jelenleg is a korpuszépítés és az elemzés fázisában van, óvakodnék egyelőre a végkövetkeztetések levonásától, s lekérdezéseimmel inkább csak a kirajzolódni látszó tendenciákat szeretném szemléltetni.

Lekérdeztem a három teljes szöveget figyelembe véve az elemzésekben szereplő összes olyan lexikai egységet, amelyre a szöveg további részében visszautalás történik (RLU), illetve az ismétlések konkrét megvalósulását (LU). Ezeket szövegenként és nyelvenként csoportosítva tartalmazza a következő táblázat (6. ábra), ahol a közel azonos magyar és német szövegenkénti mondatszám ellenére markáns eltérések körvonalazódnak.

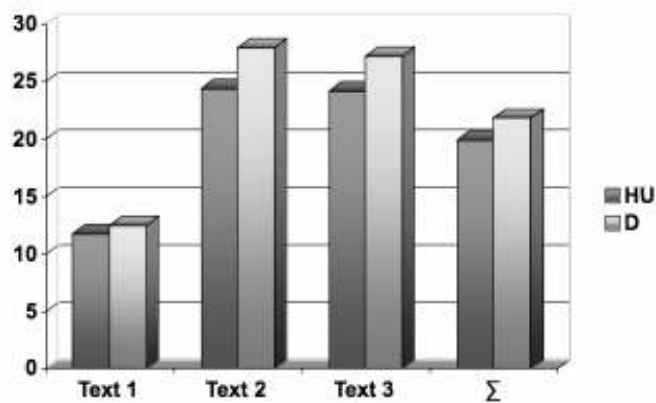
	Szöveg 1	Szöveg 2	Szöveg 3
Magyar			
Mondatszám	15	11	17
RLU	51	49	62
LU	177	268	412
Német			
Mondatszám	17	11	17
RLU	59	58	66
LU	213	308	462

1. táblázat A lexikai ismétlések realizációja

A táblából kiolvasható, hogy e három szövegből kinyert lekérdezés tapasztalata alapján a fordított német szövegek esetében nagyobb a lexikai elemek ismétlésének a sűrűsége, mint amilyen az eredeti magyar szövegekre jellemző. Ezt két grafikus ábra segítségével is szemléletesebbé teszem (7. és 8. ábra), melyekben az ismételt elemeket, illetve az ismétlések előfordulását jelenítem meg szövegenként a két nyelv viszonylatában.

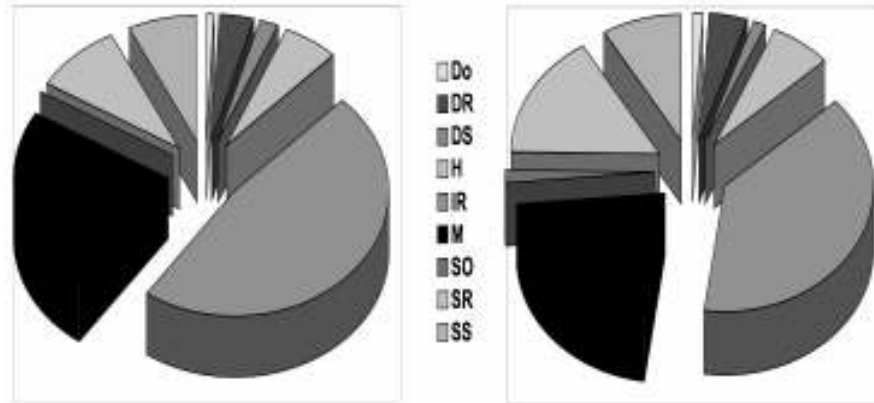


6. ábra Az ismételt elemek (RLU) sűrűsége szövegenként és összesítve



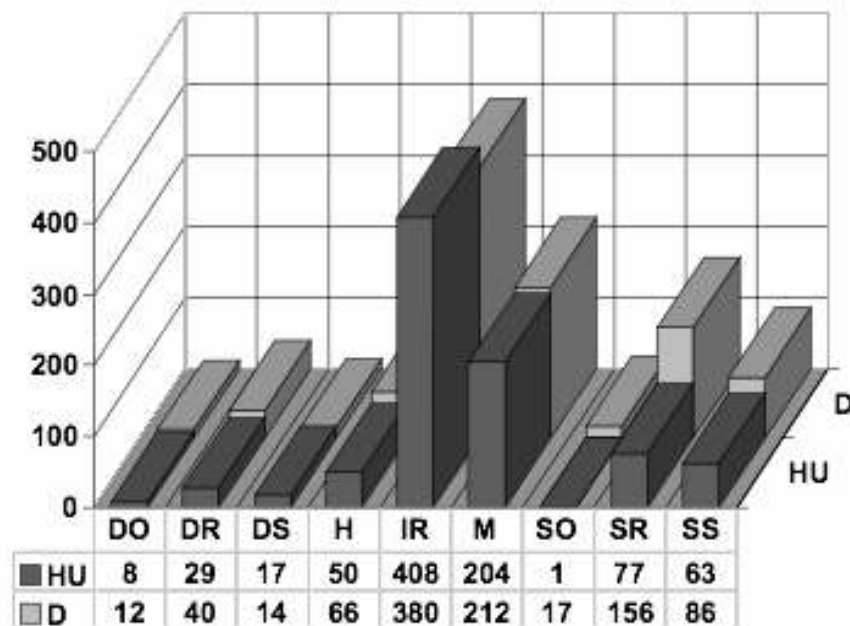
7. ábra Az ismétlések (LU) sűrűsége szövegenként és összesítve

A most bemutatott eltérések nyomán tovább haladva feltételezhető, hogy e sűrűségbeli differenciáltság együtt jár a lexikai kohéziós eszközök disztribúciójában is tapasztalható különbséggel a két vizsgált nyelv vonatkozásában. Az eloszlási arányok szemléletes megjelenítését mutatja be a körcikkelyes felosztás (9. ábra) (rövidítések: DO: derived opposites, DR: derived repetition, DS: derived synonymy, H: hyponymy, IR: instantial relations, M: meronymy, SO: simple opposites, SR: simple repetition, SS: simple synonymy).



8. ábra A lexikai elemek nyelvenkénti (magyar-német) eloszlásának aránya

Egy részletesebb elmélyedés lehetőségét kínálja a lexikai ismétlések eloszlásának arányát tartalmazó táblázat és grafikon (10. ábra) (rövidítések: vö. 9. ábra).



9. ábra A lexikai ismétlések eloszlásának aránya

Bár a lekérdezések eredményeit a rendelkezésre álló adatállomány kicsiny volta miatt egyelőre igen óvatosan kell kezelni, a kutatás jelenlegi fázisában kirajzolódó tendenciák mégis biztató előjelnek bizonyulnak a felállított hipotézisre keresett megerősítés szempontjából. Az itt közzétett ábra alapján a jelenlegi feldolgozottság szintjén úgy tűnik, hogy a vizsgált német fordított szövegek előnyben részesítik az ismétlés, az ellentét és a szinonima esetében az egyszerű alakokat a magyar forrásnyelvi szövegekre inkább jellemző derivált alakokkal szemben.

A táblázat természetesen más következtetések levonására is kínál adatokat, de ezekre szívesebben reflektálok majd nagyobb szövegállomány feldolgozása után, továbbá a tanulmány terjedelmi keretei nem teszik lehetővé a lekérdezések mondat és szószintű eredményeinek ismertetését. Mint ahogy azt a gépi elemzés előnyeinel ismertettem a program számos további lekérdezési lehetőséget tartogat, amelyek segítségével a kutatás későbbi szakaszaiban az egyes eltérések mögött húzódó ok-okozati összefüggések finomabb megvilágítására is lehetőség nyílik.

6 Összegzés

A tanulmány tartalmi mondanivalójának összefoglalásaképpen megállapíthatjuk, hogy az itt bemutatott kutatás a magyar forrásnyelvi és a fordításaként keletkezett német célnyelvi szövegekre jellemző lexikai kohéziós mintázatot kívánja egybevetni. Ezzel az összehasonlító elemzéssel ugyanakkor a szövegnyelvészet és a fordítástudomány közös kutatási körébe tartozó terület a két diszciplína összekapcsolásának új megközelítését eredményezte.

Az új megközelítést, azaz az adott nyelvpárok közötti összehasonlító elemzést a napjainkban egyre népszerűbb korpusznyelvészeti módszer alkalmazása tette lehetővé és beláthatóvá, melynek segítségével lehetővé válik az eredeti forrásnyelvi és a megfelelő célnyelvi lexikai egységek, mondatok, és szövegrészek párhuzamosítása.

A kutatás jelenlegi szakaszában kirajzolódó tendenciák tanúsága szerint a fordított német szövegekben nő a kohézióért felelős elemek sűrűsége, illetve a német szövegek az ismétlés, a szinonima és az ellentét esetében az egyszerű alakokat részesítik előnyben a magyar szövegekben népszerűbb derivált alakokkal szemben.

Bár a korpusz nagysága még korántsem biztosítja a magyar forrásnyelvű és német célnyelvű informatív szövegek kohéziós mintázatának eltéréseit kirajzoló teljes képet, de a feldolgozott szövegmennyiségből kinyert információ mindenképpen megerősíti a 2.3. pontban vázolt másik két alkorpusz létrehozásának szükségességét. Nevezetesen azokat a jövőben tervezett összehasonlító elemzéseket, amelyek tárgya a német forrásnyelvű és magyar célnyelvű, illetve az egy nyelven belül keletkezett autentikus és fordított szövegek egybevetése lesz.

Források

Művészetek Palotája. http://www.mupa.hu/epulet_ahaz.jsp?language=hu&languageCode=0
Palast der Künste. http://www.mupa.hu/epulet_ahaz.jsp?language=ge&languageCode=2
Magyar Fordítóház. <http://www.c3.hu/~bfordhaz/>

Das Ungarische Übersetzerhaus. <http://www.c3.hu/~bfordhaz/>
Balatoni hajózás. <http://www.balatonihajozas.hu/hu/rolunk.htm>
Balaton Schiffahrts AG. <http://www.balatonihajozas.hu/de/rolunk.htm>

Irodalom

- Beaugrande, R. d., Dressler W. 2000. *Bevezetés a szövegnyelvészetbe*. Budapest: Corvina.
- Balaskó, M. 2005. *Korpusznyelvészeti vizsgálatok és forrásnyelvi minták (angol és magyar tudományos szövegek anyaga alapján)*. Pécsi Tudományegyetem: PhD dolgozat.
- Catford, J.C. 1965. *A Linguistic Theory of Translation*. Oxford: Oxford University Press.
- Dróth, J. 2002. Az értékelés mint rendszeres visszajelzés a fordítás oktatásában. *Porta Lingua* 2002. *Szaknyelvoktatásunk az EU kapujában*. Debrecen: Debreceni Egyetem. 94-100.
- Halliday, M. A. K., Hasan R. 1976. *Cohesion in English*. Longman: London.
- Heltai, P. 2004. A fordító és anyelvi normák I. *Magyar Nyelvőr* Vol. 128. No. 4. 407-434.
- Heltai, P. 2005a. A fordító és anyelvi normák II. *Magyar Nyelvőr* Vol. 129. No. 1. 30-58.
- Heltai, P. 2005b. A fordító és anyelvi normák III. *Magyar Nyelvőr* Vol. 129. No. 2. 165-172.
- Károly, K. 2002. *Lexical repetition in text. A study of the text-organizing function of lexical repetition in foreign language argumentative discourse*. Frankfurt am Main: Peter Lang.
- Klaudy, K. 1987. Fordítás és aktuális tagolás. *Nyelvtudományi értekezések* 123. Budapest: Akadémiai Kiadó.
- Klaudy, K. 2004. A kommunikatív szakaszhatárok eltűnése a magyarra fordított európai uniós szövegekben. *Magyar Nyelvőr* Vol. 128. No. 4. 389-407.
- Nagy, Szikszainé I. 1999. *Leíró magyar szövegtan*. Budapest: Osiris.
- Prószéky, G., Kis B. 1999. *Számítógéppel emberi nyelven. Intelligens szövegkezelés számítógéppel*. Bicske: SZAK Kiadó.
- Szabó Mihály, G. 2003. A Szlovákiai magyar szakfordítások minőségének javításáról és az objektív fordításkritika megteremtésének feltételeiről. *Társadalomtudományi Szemle* Vol. V. No. 4. 55-68.
- Szirmai, M. 2005. *Bevezetés a korpusznyelvészetbe. A korpusznyelvészet alkalmazása az anyanyelv és az idegen nyelv tanulásában és tanításában*. Budapest: Tinta Könyvkiadó.
- Tolcsvai Nagy, G. 1994. *A szövegtan kialakulása*. Elérhető:
<http://www.jgytf.u-szeged.hu/~vass/szemml112.htm>
- Vass, L. 2002. *Előadásvázlatok. Szövegtan*. A szenzus. A rendszerszerű szenzus. Diktum, apperceptum, evokátum. Kohézió. Kohézióhordozók. Elérhető:
<http://www.jgytf.u-szeged.hu/~vass/vsze111.htm>